Review 115A : Inner products and norms

From now on, F is either $\mathbb{R}$ or $\mathbb{C}$.

Def Let V be a v.s. over F. An inner product on V is a function $(x,y) \mapsto \langle x,y\rangle$ from V^2 to F s.t.

a) $\langle x+z,y\rangle = \langle x,y\rangle + \langle z,y\rangle \quad \forall x,y,z \in V$

b) $\langle cx,y\rangle = c\langle x,y\rangle \quad \forall x,y \in V, c \in F.$

c) $\overline{\langle x,y\rangle} = \langle y,x\rangle$ (complex conjugate), $\forall x,y \in V$.

d) $\langle x,x\rangle > 0$ if $x \neq 0 \quad \forall x \in V$.

Ex 1 The standard inner product on F^n is given by: for $x=(a_1,\dots,a_n)$, $y=(b_1,\dots,b_n)$ in F^n,

$\langle x,y\rangle = \sum_{i=1}^{n} a_i \overline{b_i}$.

Def For $A \in M_{m\times n}(F)$, its conjugate transpose (or adjoint) is the $n\times m$ matrix A^* s.t. $A^*_{ij} = \overline{A_{ji}} \quad \forall i,j$.

Def A v.s. V over F endowed with a specific inner product is called an inner product space.
(Note: there can be many different inner products on the same v.s.)
If $F=\mathbb{C}$ — complex inner prod. space.
If $F=\mathbb{R}$ — real inner prod. space.

Thm 6.1 Let V be an I.P.S. (Inner Prod. Space). For all $x,y,z \in V$ and $c \in F$:

a) $\langle x,y+z\rangle = \langle x,y\rangle + \langle x,z\rangle$.

b) $\langle x,cy\rangle = \bar{c}\langle x,y\rangle$.

c) $\langle x,0\rangle = \langle 0,x\rangle = 0$.

d) $\langle x,x\rangle = 0$ iff $x=0$.

e) If $\langle x,y\rangle = \langle x,z\rangle$ for all $x \in V$, then $y=z$.

Def Let V be an I.P.S. For $x \in V$, the norm (or length) of x is $\|x\| = \sqrt{\langle x,x\rangle}$.

Ex Let $V=\mathbb{R}^n$ with the standard inner prod. If $x=(a_1,\dots,a_n)$, then
$\|x\| = \|(a_1,\dots,a_n)\| = \left(\sum_{i=1}^{n} |a_i|^2\right)^{\frac{1}{2}}$ is the usual Euclidean length of a vector. If $n=1$, $\|a\| = |a|$.

Thm 6.2 Let V be an I.P.S. over F. Then $\forall x,y \in V$ and $c \in F$:

a) $\|cx\| = |c|\cdot\|x\|$.

b) $\|x\|=0$ iff $x=0$; always $\|x\| \geq 0$.

c) (Cauchy-Schwarz inequality) $|\langle x,y\rangle| \leq \|x\|\cdot\|y\|$.

d) (Triangle inequality) $\|x+y\| \leq \|x\| + \|y\|$.

Def Let V be an I.P.S.
Vectors $x,y \in V$ are orthogonal if $\langle x,y\rangle = 0$.
A set $S \subseteq V$ is orthogonal if any two distinct vectors in S are orthogonal.
A vector $x \in V$ is a unit vector if $\|x\|=1$.
A set $S \subseteq V$ is orthonormal if S is orthogonal and consists of unit vectors.
Note: for any $x \in V$, $\left(\frac{1}{\|x\|}\right)\cdot x$ is a unit vector.

115A : Gram-Schmidt orthogonalization and orthogonal complements.

Def Let V be an I.P.S. A subset of V is an orthonormal basis if it is an ordered basis that is orthonormal.

Ex The standard basis for F^n is orthonormal.

Thm 6.3 Let V be an I.P.S. and $S=\{v_1,\dots,v_k\}$ an orthogonal subset of V with $v_i \neq 0$.
If $y \in Span(S)$, then $y = \sum_{i=1}^{k} \frac{\langle y,v_i\rangle}{\|v_i\|^2} v_i$.

Cor 1 If in addition S is orthonormal and $y \in Span(S)$, then
$y = \sum_{i=1}^{k} \langle y,v_i\rangle v_i$.

Cor 2 Let V be an I.P.S. and $S \subseteq V$ orthogonal, $v \neq 0$ for all $v \in S$.
Then S is lin. indep.

Thm 6.4. (Gram - Schmidt orthogonalization) Let V be an I.P.S. and $S = \{w_1, \dots, w_n\} \subseteq V$ lin. indep.
Let $S' = \{v_1, \dots, v_n\}$, where $v_1 = w_1$ and v_k are defined recursively as follows:
$v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j \rangle}{\|v_j\|^2} v_j$ for $2 \leq k \leq n$.

Then S' is an orthogonal set of non-zero vectors and $Span(S') = Span(S)$.

Using it, one obtains:
Thm 6.5 Let $V \neq \{0\}$ be an I.P.S. with $\dim(V) < \infty$.
Then V has an orthonormal basis $\beta = \{v_1, \dots, v_n\}$.
And for any $x \in V$, $x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i$.

Cor Let V be an I.P.S. with $\dim(V) < \infty$ and an orthonormal basis $\beta = \{v_1, \dots, v_n\}$. Let $T \in \mathcal{L}(V)$, $A = [T]_\beta$.
Then for any i, j we have $A_{ij} = \langle T(v_j), v_i \rangle$.

Def Let $\emptyset \neq S \subseteq V$ and V an I.P.S. Let
$S^\perp = \{x \in V : \langle x, y \rangle = 0$ for all $y \in S\}$ – the orthogonal complement of S.
Note: $S^\perp$ is a subspace of V for any $S \subseteq V$, $\{0\}^\perp = V$ and $V^\perp = 0$.

Thm 6.6 Let W with $\dim(W) < \infty$ be a subspace of an I.P.S. V. Let $y \in V$.
Then there exist unique vectors $u \in W$ and $z \in W^\perp$ s.t. $y = u + z$.
Moreover, if $\{v_1, \dots, v_k\}$ is an orthonormal basis for W, then
$u = \sum_{i=1}^{k} \langle y, v_i \rangle v_i$.

Def The vector u in the theorem is called the orthogonal projection of y on W.
Cor u is the unique vector in W that is "closest" to y: for any $x \in W$, $\|y - x\| \geq \|y - u\|$, and this is an equality iff $x = u$.

Thm 6.7 Supp. $S = \{v_1, \dots, v_k\}$ is an orthonormal set in an n-dim. inner product space V. Then:
a) S can be extended to an orthonormal basis $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ for V.
b) If $W = Span(S)$, then $S_1 = \{v_{k+1}, \dots, v_n\}$ is an orthonormal basis for $W^\perp$.
c) If W is any subspace of V, then $\dim(V) = \dim(W) + \dim(W^\perp)$, and $V = W \oplus W^\perp$.

115A: The adjoint of a lin. operator.
Let V be an I.P.S., and $y \in V$.
Then the function $g : V \to F$ defined by $g(x) = \langle x, y \rangle$ is linear. Conversely:
Thm 6.8. Let V be an I.P.S. over F, $\dim(V) < \infty$.
Let $g : V \to F$ be a lin. transformation.
Then there exists a unique vector $y \in V$ s.t. $g(x) = \langle x, y \rangle$ for all $x \in V$.

Thm 6.9 Let V be a fin. dim. I.P.S., and let $T \in \mathcal{L}(V)$.
Then there exists a unique function $T^* : V \to V$ s.t.
$\langle T(x), y \rangle = \langle x, T^*(y) \rangle$ for all $x, y \in V$.
Moreover, T^* is linear.

Def The unique lin. op. T^* on V satisfying $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$ for all $x, y \in V$ is called the adjoint of T.

Remark Every $T \in \mathcal{L}(V)$ has an adjoint if $\dim(V) < \infty$. This is not true if $\dim(V) = \infty$!
(see Textbook, Exercise 6.3.24)

Thm 6.10 Let V be an I.P.S. with $\dim(V) < \infty$, let β be an orthonormal basis for V.
If $T \in \mathcal{L}(V)$, then $[T^*]_\beta = [T]_\beta^*$.

Cor If $A \in M_{n\times n}(F)$, then $L_{A^*} = (L_A)^*$.

Thm 6.11 Let V be an I.P.S. and $T, U \in \mathcal{L}(V)$. Then:
a) $(T+U)^* = T^* + U^*$.
b) $(cT)^* = \bar{c}T^*$ for any $c \in F$.
c) $(TU)^* = U^*T^*$.
d) $T^{**} = T$.
e) $I^* = I$.

(the same hold for matrices and their conjugate transposes).

Normal and self-adjoint operators

Given $T \in \mathcal{L}(V)$, we know:
V has a basis consisting of eigenvectors for T $\Leftrightarrow$ T is diag'z.
Now want to understand:
V has an orthonormal basis of eigenvectors for T $\Leftrightarrow$ T is ???

Lemma Let $T \in \mathcal{L}(V)$, V an I.P.S. with $\dim(V) < \infty$.
If T has an e.vect., then so does T^*.

Proof Supp. $v \neq 0$ is an e.vect of T with e.val. λ.
Then for any $x \in V$:
$0 = \langle 0, x\rangle = \langle (T-\lambda I)(v), x\rangle = \langle v, (T-\lambda I)^*(x)\rangle \overset{\text{Thm 6.11}}{=} \langle v, (T^* - \bar{\lambda} I)(x)\rangle$
hence v is orthogonal to $R(T^* - \bar{\lambda}I)$.
In particular, $v \notin R(T^* - \bar{\lambda}I)$ as $v \neq 0$, so the lin. op. $T^* - \bar{\lambda}I$ is not onto.
As $\dim(V) < \infty$, this implies it is also not one-to-one, so $N(T^* - \bar{\lambda}I) \neq \{0\}$.
Take any $0 \neq w \in N(T^* - \bar{\lambda}I)$, then $T^*(w) = \bar{\lambda}w$, hence w is an e.vect. of T^* with e.val. $\bar{\lambda}$.

Thm 6.14 (Schur) Let $T \in \mathcal{L}(V)$ for V an I.P.S. with $\dim(V) < \infty$.
Supp. that the char. poly. of T splits.
Then there exists an orthonormal basis β for V s.t. the matrix $[T]_\beta$ is upper triangular.

Proof.
By induction on $n = \dim(V)$.
Clear for $n=1$ (as any 1×1 matrix is upper triangular).
Suppose holds for $n-1$.
Let $T \in \mathcal{L}(V)$ with $\dim(V) = n$ be given, and its char. poly. $f(t)$ splits.
Then T has an e.vect. (recall by Thm 5.2: $\lambda \in F$ is an e.val of T iff $f(\lambda) = 0$. As $f(t)$ splits over F, it has a root in F — hence some e.val. $\lambda \in F$, and there is an e.vect. $v \in V$ corresp. to it).
By the previous lemma, T^* also has an e.vect. z, and normalizing it we may assume z is a unit e.vect. with e.val. λ.
Let $W = \mathrm{Span}(\{z\})$.

Claim $W^\perp$ is T-invariant.
If $y \in W^\perp$ and $x \in W$, then $x = cz$ for some $c \in F$, hence
$\langle T(y), x\rangle = \langle T(y), cz\rangle = \langle y, T^*(cz)\rangle = \langle y, cT^*(z)\rangle = \langle y, c\lambda z\rangle = \overline{c\lambda}\langle y, z\rangle \overset{\text{as } y \in W^\perp, \text{ so orthogonal to } z}{=} \overline{c\lambda}\cdot 0 = 0.$
So $T(y) \in W^\perp$. Hence $W^\perp$ is T-inv.

By Thm 5.21, the char. poly. of $T_{W^\perp}$ divides the char. poly. of T, so it also splits.
By Thm 6.7 (c), $\dim(W^\perp) = \dim(V) - \dim(W) = \dim(V) - 1 = n-1$.
By the induction hypothesis applied to $T_{W^\perp}$ we obtain an orthonormal basis γ for $W^\perp$ s.t. $[T_{W^\perp}]_\gamma$ is upper triangular.
Then $\beta = \gamma \cup \{z\}$ is an orthonormal set (as $z \in W$ and $\gamma \subseteq W^\perp$ is orthogonal, and z is a unit ~~vector~~).
Then β is lin. indep. (by Cor 2 of Thm 6.3) of size $n = \dim(V)$, so β is an orthonormal basis for V. And, if $\gamma = \{z_1, \dots, z_{n-1}\}$, we have:

$$[T]_\beta = \begin{pmatrix} | & & | & | \\ [T(z_1)]_\beta & \cdots & [T(z_{n-1})]_\beta & [T(z)]_\beta \\ | & & | & | \end{pmatrix} = \begin{pmatrix} | & & | & | \\ [T(z_1)]_\gamma & \cdots & [T(z_{n-1})]_\gamma & [T(z)]_\beta \\ 0 & & 0 & | \end{pmatrix} = \left(\begin{array}{c|c} [T_{W^\perp}]_\gamma & [T(z)]_\beta \\ \hline 0 \cdots 0 & | \end{array}\right),$$

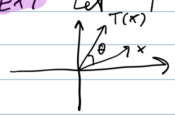
and since $[T_{W^\perp}]_\gamma$ is upper-triangular, this matrix $[T]_\beta$ is also upper triangular.

Now we return to determining what ??? above should be.
Assume that V admits an orthonormal basis β of e.vect's for T.
Then in particular $[T]_\beta$ is a diagonal matrix.
But then $[T^*]_\beta = [T]_\beta^*$ is also diagonal.
As diagonal matrices commute, we conclude that T and T^* commute. This motivates:

Def Let V be an I.P.S., and $T \in \mathcal{L}(V)$. We say T is normal if $TT^* = T^*T$.
An $n \times n$ real or complex matrix A is normal if $AA^* = A^*A$.

So V admits an orthonormal basis of e.vect's for $T \Rightarrow T$ is normal.
What about "$\Leftarrow$" ?

Ex 1 Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be rotation by θ, $0 < \theta < \pi$. Let β be the standard basis for $\mathbb{R}^2$.



Then $[T]_\beta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Calculating, we get $AA^* = I_2 = A^*A$, so T is normal.
But T has no e.vect's at all! (every vector gets rotated, so can't be just scaled).

This shows that "$\Leftarrow$" can fail for $F = \mathbb{R}$. We show that at least "$\Leftarrow$" holds for $F = \mathbb{C}$!

Thm 6.15 Let V be an I.P.S., let $T \in \mathcal{L}(V)$ be normal. Then:
a) $\|T(x)\| = \|T^*(x)\| \quad \forall x \in V$.
b) $T - cI$ is normal $\forall c \in F$.
c) If x is an e.vect. of T, then x is also an e.vect. of T^*.
In fact, if $T(x) = \lambda x$, then $T^*(x) = \bar{\lambda} x$.
d) If $\lambda_1 \neq \lambda_2$ are e.vals of T with corresp. e.vect's x_1, x_2, then x_1 and x_2 are orthogonal.

Proof. a) For any $x \in V$,
$$\|T(x)\|^2 = \langle T(x), T(x)\rangle \underset{\text{Thm 6.11(d)}}{=} \langle T(x), (T^*)^*(x)\rangle \underset{\text{def. of adjoint op.}}{=} \langle T^*T(x), x\rangle \underset{\text{as } T \text{ is normal}}{=} \langle TT^*(x), x\rangle \underset{\text{def. of adjoint op.}}{=} \langle T^*(x), T^*(x)\rangle = \|T^*(x)\|^2.$$
Hence $\|T(x)\| = \|T^*(x)\|$ as norm is always non-negative.
b) By Thm 6.11,
$$(T - cI)^* = T^* - \bar{c}I.$$
Then $(T-cI)(T-cI)^* = (T-cI)(T^* - \bar{c}I) = TT^* - T\bar{c}I - cIT^* + cI\bar{c}I =$
$$= \underset{\text{as } T \text{ is normal}}{T^*T} - \underset{\text{as } T \text{ is lin.}}{\bar{c}IT} - \underset{\text{as } T^* \text{ is lin.}}{T^*cI} + \bar{c}IcI.$$
And $(T-cI)^*(T-cI) = (T^* - \bar{c}I)(T-cI) = T^*T - T^*cI - \bar{c}IT + \bar{c}IcI$

Hence $(T-cI)(T-cI)^* = (T-cI)^*(T-cI)$, so $T - cI$ is normal.

c) Supp. $T(x) = \lambda x$ for some $x \in V$, $\lambda \in F$. Let $U = T - \lambda I$.
Then $U(x) = 0$, and U is normal by (b). Thus (a) implies:
$$0 = \|U(x)\| = \|U^*(x)\| \underset{\text{Thm 6.11}}{=} \|(T^* - \bar{\lambda} I)(x)\| = \|T^*(x) - \bar{\lambda} x\|. \quad \text{Hence } T^*(x) = \bar{\lambda} x.$$

d) Let $\lambda_1 \neq \lambda_2$ be e.val's of T with corresp. e.vect's x_1, x_2. By (c):
$$\lambda_1 \langle x_1, x_2 \rangle = \langle \lambda_1 x_1, x_2 \rangle = \langle T(x_1), x_2 \rangle = \langle x_1, T^*(x_2) \rangle = \langle x_1, \bar{\lambda}_2 x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle.$$
Since $\lambda_1 \neq \lambda_2$, this implies $\langle x_1, x_2 \rangle = 0$.

Thm 6.16 Let $T \in \mathcal{L}(V)$, V a C.I.P.S. (Complex I.P.S.) with $\dim(V) < \infty$.
Then T is normal $\Leftrightarrow$ there exists an orthonormal basis for V consisting of e.vect's for T.

Proof $\Rightarrow$ Supp. T is normal.
By the fundamental theorem of algebra, every polynomial over $F = \mathbb{C}$ splits.
In particular, the char. poly. of T splits.
By Shur's thm 6.14, there is an orthonormal basis $\beta = \{v_1, \dots, v_n\}$ for V s.t. $[T]_\beta = A$ is upper triangular.
Then v_1 is an e.vect. of T (as $[T]_\beta$ is upper triang, $[T(v_1)]_\beta = \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ for some $\lambda \in F$, so $T(v_1) = \lambda v_1$).

Assume that $v_1, \dots, v_{k-1}$ are e.vect's of T. We claim that then v_k is also an e.vect. of T.
(as then by induction, starting with v_1, all v_i are e.vect's of T and we are done).
Consider any $j < k$.
Let $\lambda_j \in F$ denote the e.val. of T corresponding to v_j.
By Thm 6.15, $T^*(v_j) = \bar{\lambda}_j v_j$.
We have:
$$[T(v_k)]_\beta = A \cdot [v_k]_\beta = A \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \}k-1 \\ \leftarrow k \\ \}n-k \end{matrix} = \begin{pmatrix} A_{11} \cdot 0 + \dots + A_{1,k-1} \cdot 0 + A_{1,k} \cdot 1 + A_{1,k+1} \cdot 0 + \dots + A_{1,n} \cdot 0 \\ \vdots \\ A_{n1} \cdot 0 + \dots + A_{n,k-1} \cdot 0 + A_{n,k} \cdot 1 + A_{n,k+1} \cdot 0 + \dots + A_{n,n} \cdot 0 \end{pmatrix} =$$
$$= \begin{pmatrix} A_{1,k} \\ \vdots \\ A_{n,k} \end{pmatrix} = \begin{pmatrix} A_{1k} \\ \vdots \\ A_{kk} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{as } A_{i,k} = 0 \text{ for all } i > k \text{ since } A \text{ is upper triangular.}$$

That is, $T(v_k) = A_{1k} \cdot v_1 + \dots + A_{kk} \cdot v_k$.
And by Corollary to Thm 6.5 we have:
$$A_{jk} = \langle T(v_k), v_j \rangle = \langle v_k, T^*(v_j) \rangle = \langle v_k, \bar{\lambda}_j v_j \rangle = \lambda_j \langle v_k, v_j \rangle \overset{\text{as } \beta \text{ is orthogonal, and } j<k}{=} 0.$$
As this holds for every $j < k$, we conclude
$T(v_k) = A_{kk} v_k$, hence v_k is an e.vect. of T.

$\Leftarrow$ Already proved just before the definition of "normal".

What about R.I.P.S. (Real I.P.S.)?
Example 1 shows that "T normal" is not enough to guarantee existence of an orthonormal basis of e.vect's. So we need a stronger condition:

Def Let $T \in \mathcal{L}(V)$ for V an I.P.S.
T is self-adjoint (Hermitian) if $T = T^*$.
A matrix $A \in M_{n \times n}(F)$ is self-adjoint if $A = A^*$.

Remark Of course, T self-adjoint $\Rightarrow$ T is normal, and
T is self-adjoint $\Leftrightarrow$ $[T]_\beta$ is self-adjoint for an orthonormal basis β.

Lemma Let $T \in \mathcal{L}(V)$ be self-adjoint and $\dim(V) < \infty$. Then:

a) Every e.val. of T is real
(holds by definition if $F = \mathbb{R}$, but is meaningful for $F = \mathbb{C}$).

b) Supp. V is a R.I.P.S. Then the char. poly of T splits.

Proof

a) Supp. $T(x) = \lambda x$ for $x \neq 0$ in V.
As a self-adjoint operator is also normal, by Thm 6.15(c) we have:
$\lambda x = T(x) \underset{\text{self-adjoint}}{=} T^*(x) = \bar{\lambda} x$. As $x \neq 0$, this implies $\bar{\lambda} = \lambda$, hence λ is real.

b) Let $n = \dim(V)$, β an orthonormal basis for V, and $A = [T]_\beta$.
Then A is self-adjoint.
Let T_A be the lin. op. on $\mathbb{C}^n$ defined by $T_A(x) = Ax$ for all $x \in \mathbb{C}^n$. (note that $F = \mathbb{R}$).
Note that T_A is self-adjoint because $[T_A]_\gamma = A$, where γ is the standard ordered (orthonormal) basis for $\mathbb{C}^n$.
By (a), the e.val's of T_A are real.
By the fundamental theorem of algebra, the char. poly of T_A splits into factors of the form $t - \lambda$, for λ an e.val. of T_A.
Since each e.val. λ of T_A is real, i.e. $\lambda \in \mathbb{R}$, it follows that the char. poly. of T_A already splits over $\mathbb{R}$!
But T_A has the same char. poly as A, which has the same char. poly. as T.
Therefore the char. poly. of T splits.

We are ready to prove an analog of Thm 6.16 for R.I.P.S. instead of C.I.P.S.

Thm 6.17 Let $T \in \mathcal{L}(V)$ for V a R.I.P.S. with $\dim(V) < \infty$.
Then T is self-adjoint $\Leftrightarrow$ there exists an orthonormal basis β for V consisting of e.vect's of T.

Proof. $\Rightarrow$

Supp. T is self-adjoint.
By the lemma, (b), the char. poly. of T splits over $F = \mathbb{R}$.
Applying Shur's Thm 6.14, we find an orthonormal basis β for V s.t. the matrix $A = [T]_\beta$ is upper triangular. But
$A^* = [T]_\beta^* = [T^*]_\beta = [T]_\beta = A$ as T is self-adjoint.
So A and A^* are both upper triangular, therefore A must be a diagonal matrix.
Thus β must consist of e.vect's of T.

$\Leftarrow$ If β is an orthonormal basis for V consisting of e.vect's for T, then we see that

$$[T]_\beta = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \quad \text{for some } \lambda_i \in F = \mathbb{R}.$$

But $[T^*]_\beta = [T]_\beta^*$ by Thm 6.10, and since $\lambda_i \in \mathbb{R}$ we have $\bar{\lambda}_i = \lambda_i$ so
$[T^*]_\beta = [T]_\beta$, hence $T^* = T$, so T is normal.

Unitary and orthogonal operators and their matrices

In this section we study lin. operators that preserve length of vectors.

Def Let $T \in \mathcal{L}(V)$ for V an I.P.S. over F with $\dim(V) < \infty$.
T is an isometry if $\|T(x)\| = \|x\|$ for all $x \in V$.
If $F = \mathbb{C}$, we call such T unitary.
If $F = \mathbb{R}$, we call such T orthogonal.

Ex Rotation and reflection operators on $\mathbb{R}^2$ are orthogonal.

Thm 6.18 Let $T \in \mathcal{L}(V)$, V an I.P.S. with $\dim(V) < \infty$. The following are equivalent:
a) $TT^* = T^*T = I$. (hence every isometry is a normal operator, but not vice versa!)
b) $\langle T(x), T(y)\rangle = \langle x, y\rangle$ for all $x, y \in V$.
c) If β is an orthonormal basis for V, then $T(\beta)$ is an orthonormal basis for V.
d) There exists an orthonormal basis for V s.t. $T(\beta)$ is an orthonormal basis for V.
e) $\|T(x)\| = \|x\|$ for all $x \in V$, that is T is an isometry.

To prove it, we will need the following

Lemma Let U be a self-adjoint op. on a fin. dim I.P.S. V.
If $\langle x, U(x)\rangle = 0$ for all $x \in V$, then $U = T_0$ – the zero-transformation on V.

Proof By Thm 6.16 if $F = \mathbb{C}$ or Thm 6.17 if $F = \mathbb{R}$, there is an orthonormal basis β for V consisting of e.vect's for U.
If $x \in \beta$, then $U(x) = \lambda x$ for some $\lambda \in F$. Thus
$0 = \langle x, U(x)\rangle = \langle x, \lambda x\rangle = \bar{\lambda}\langle x, x\rangle$.
As $x \in \beta$ is $\neq 0$, $\langle x, x\rangle > 0$. Hence $\bar{\lambda} = 0$.
So $U(x) = 0$ for all $x \in \beta$, thus as U is linear, $U(x) = 0$ for all $x \in V$, so $U = T_0$.

Proof of Thm 6.18

(a) => (b) Let $x, y \in V$.
Then $\langle x, y\rangle = \langle \underbrace{T^*T}_{=I \text{ by (a)}}(x), y\rangle \underset{\text{adjoint of } T^*}{=} \langle T(x), (T^*)^*(y)\rangle \underset{\text{Thm 6.11 (d)}}{=} \langle T(x), T(y)\rangle$.

(b) => (c) Let $\beta = \{v_1, \dots, v_n\}$ be an orthonormal basis for V.
So $T(\beta) = \{T(v_1), \dots, T(v_n)\}$.
As $\langle v_i, v_j\rangle = \delta_{ij}$, by (b) we have $\langle T(v_i), T(v_j)\rangle = \delta_{ij}$.
And $\|T(v_i)\|^2 = \langle T(v_i), T(v_i)\rangle \overset{(b)}{=} \langle v_i, v_i\rangle = \|v_i\|^2 = 1$, hence all $T(v_i)$ are unit vectors, and $T(\beta)$ is an orthonormal basis for V.

(c) => (d) Obvious, as by (c), (d) holds for any orthonormal β.

(d) => (e) Let $x \in V$, and $\beta = \{v_1, \dots, v_n\}$. Then
$x = \sum_{i=1}^{n} a_i v_i$ for some $a_i \in F$. So:

$$\|x\|^2 = \left\langle \sum_{i=1}^{n} a_i v_i, \sum_{j=1}^{n} a_j v_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i \bar{a}_j \langle v_i, v_j\rangle \overset{\beta \text{ is orthogonal}}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} a_i \bar{a}_j \delta_{ij} = \sum_{i=1}^{n} a_i \bar{a}_i = \sum_{i=1}^{n} |a_i|^2.$$

Applying the same calculation to $T(x) = \sum_{i=1}^{n} a_i T(v_i)$ and using that $T(\beta)$ is orthonormal by (d), we get
$\|T(x)\|^2 = \sum_{i=1}^{n} |a_i|^2$.
Hence $\|T(x)\| = \|x\|$.

(e) => (a) For any $x \in V$ we have
$\langle x, x\rangle = \|x\|^2 \underset{(e)}{=} \|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, T^*T(x)\rangle$.

So $\langle x, (I - T^*T)(x)\rangle = 0$ for all $x \in V$.
Let $U = I - T^*T$. Then U is self-adjoint ($U^* = (I - T^*T)^* \overset{\text{Thm 6.11}}{=} I^* - (T^*T)^* = I - T^*T^{**} = I - T^*T = U$).
And $\langle x, U(x)\rangle = 0$ for all $x \in V$. Then by the Lemma $T_0 = U = I - T^*T$, therefore $T^*T = I$.
Recall that for two square matrices $A, B \in M_{n\times n}(F)$, if $AB = I$, then also $BA = I$ (115A).

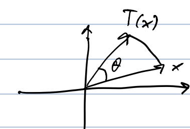
Hence, taking β any orthonormal basis, $T^*T = I \Rightarrow [T^*]_\beta [T]_\beta = I \Rightarrow [T]_\beta [T^*_\beta] = I$
$\Rightarrow TT^* = I$, so (a) holds.

Cor 1 Let $T \in \mathcal{L}(V)$ on V a R.I.P.S. with $\dim(V) < \infty$. Then:
V has an orthonormal basis of e.vect's of T with the corresp. e.val's of absolute value 1
$\Leftrightarrow$
T is both self-adjoint and orthogonal.

Proof
$\Rightarrow$ Supp. V has an orthonormal basis $\beta = \{v_1, \dots, v_n\}$ s.t. $T(v_i) = \lambda_i v_i$ and $|\lambda_i| = 1$ for all i.
By Thm 6.17, T is self-adjoint.
Hence $(TT^*)(v_i) = T(T^*(v_i)) = T(T(v_i)) = T(\lambda_i v_i) \overset{T \text{ lin.}}{=} \lambda_i \lambda_i v_i = \lambda_i^2 v_i = v_i$ for each i.
So $TT^* = I$. And then, as in the end of the previous proof, also $T^*T = I$, so T is orthogonal by Thm 6.18 (a)
$\Leftarrow$ If T is self-adjoint, then by Thm 6.17 V has an orthonormal basis $\beta = \{v_1, \dots, v_n\}$ s.t. $T(v_i) = \lambda_i v_i$.
If T is also orthogonal, we have
$|\lambda_i| \cdot \|v_i\| = \|\lambda_i v_i\| = \|T(v_i)\| = \|v_i\|$, so $|\lambda_i| = 1$ for all i.

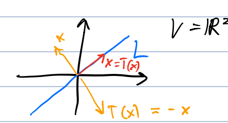
Cor 2 Let $T \in \mathcal{L}(V)$ for V a C.I.P.S. with $\dim(V) < \infty$. Then:
V has an orthonormal basis of e.vect's of T with corresp. e.val's of absolute value 1
$\Leftrightarrow$ T is unitary.
Proof Similar to the proof of Corollary 1 (see H/W 6.)

Ex 1 Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be a rotation by θ, $0 < \theta < \pi$.



Then T preserves length of vectors, hence it also preserves the standard inner product on $\mathbb{R}^2$ by Thm 6.18.
For β standard basis, $[T]_\beta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. A direct calculation shows that $[T]_\beta \neq [T]^*_\beta$, hence T is not self-adjoint.
Also follows by Thm 6.15 as there are no e.vect's.
In fact, T^* is the rotation by $-\theta$.

Def Let L be a 1-dimensional subspace of $V = \mathbb{R}^2$.



Then L is a line through the origin.
A lin. op. T on $\mathbb{R}^2$ is called a reflection of $\mathbb{R}^2$ about L
if $T(x) = x$ for all $x \in L$ and $T(x) = -x$ for all $x \in L^\perp$.

Ex 3 Let T be a reflection of $\mathbb{R}^2$ about a line L through the origin.
Then T is an orthogonal operator:
Take any $v_1 \in L$, $v_2 \in L^\perp$ with $\|v_1\| = \|v_2\| = 1$.
Then by def. $T(v_1) = v_1$ and $T(v_2) = -v_2$.
Thus v_1, v_2 are e.vect's with e.val's 1 and -1, respectively.
And $\{v_1, v_2\}$ is an orthonormal basis for $\mathbb{R}^2$.
Hence T is an orthogonal operator by Corollary 1 above.

Matrices of unitary and orthogonal operators

Def A square matrix A is orthogonal if $A^tA = AA^t = I$ and unitary if $A^*A = AA^* = I$.

As always, by Thm 6.10, T is unitary/orthogonal $\Leftrightarrow$ $[T]_\beta$ is unitary/orthogonal, for some orthonormal basis β of V.

Remark 1) $AA^* = I \iff$ the rows of A form an orthonormal basis for F^n.
As $\delta_{ij} = I_{ij} = (AA^*)_{ij} = \sum_{k=1}^{n} A_{ik} A^*_{kj} = \sum_{k=1}^{n} A_{ik} \overline{A_{jk}}$, and the last term is the (standard) inner prod. of the i^{th} and j^{th} rows of A.
2) $A^*A = I \iff$ the columns of A form an orthonormal basis for F^n.

Def We say that a matrix $A \in M_{n\times n}(F)$ is unitarily / orthogonally equivalent to a matrix $B \in M_{n\times n}(F)$ if there exists a unitary / orthogonal matrix $P \in M_{n\times n}(F)$ s.t. $A = P^*BP$.

Remark This is an equivalence relation on $M_{n\times n}(F)$ (see H/W 6).

Thm 6.19 Let $A \in M_{n\times n}(\mathbb{C})$. Then A is normal $\iff$ A is unitarily equivalent to a diagonal matrix.

Proof.

$\Leftarrow$ Suppose $A = P^*DP$ for P a unitary matrix and D a diagonal matrix. Then:

$$AA^* = (P^*DP)(P^*DP)^* \underset{\text{Thm 6.11}}{=} (P^*DP)(P^*D^*P^{**}) = (P^*DP)(P^*D^*P) = P^*DPP^*D^*P \underset{P \text{ unitary}}{=}$$
$$= P^*DID^*P = P^*DD^*P.$$

Similarly, $A^*A = P^*D^*DP$.
Since D is a diagonal matrix, $DD^* = D^*D$. Thus $AA^* = A^*A$, so A is normal.

$\Rightarrow$ Assume A is normal.
By Thm 6.16 applied to L_A, there exists an orthonormal basis β for F^n consisting of eigenvectors of A.
By Thm 2.23 (its corollary) from 115A, $[L_A]_\beta = Q^{-1}AQ$, where Q is the $n\times n$ matrix whose j^{th} column is the j^{th} vector of β; and $D = [L_A]_\beta$ is a diagonal matrix.
Hence the columns of Q form an orthonormal basis for F^n, so by the remark above $Q^*Q = I$. But since these are square matrices, then also $QQ^* = I$, so Q is unitary and $Q^* = Q^{-1}$.
Hence $D = Q^*AQ$, so A is unitarily equivalent to a diagonal matrix.

Thm 6.20 Let $A \in M_{n\times n}(\mathbb{R})$. Then A is symmetric $\iff$ A is orthogonally equivalent to a real diagonal matrix.

Proof Similar to Thm 6.19 (see H/W 6).

Finally, we have a matrix form of Schur's theorem:

Thm 6.21 Let $A \in M_{n\times n}(F)$ be such that its char. poly. splits over F.
a) If $F = \mathbb{C}$, then A is unitarily equivalent to a complex upper triangular matrix.
b) If $F = \mathbb{R}$, then A is orthogonally equivalent to a real upper triangular matrix.